Domestic Heat Demand Prediction Using Neural Networks

Vincent Bakker, Albert Molderink, Johann L. Hurink and Gerard J.M. Smit
University of Twente, Department of EEMCS
P.O. Box 217, 7500 AE Enschede, The Netherlands
v.bakker@utwente.nl

Abstract

By combining a cluster of microCHP appliances, a virtual power plant can be formed. To use such a virtual power plant, a good heat demand prediction of individual households is needed since the heat demand determines the production capacity. In this paper we present the results of using neural networks techniques to predict the heat demand of individual households. This prediction is required to determine the electricity production capacity of the large fleet of microCHP appliances. All predictions are short-term (for one day) and use historical heat demand and weather influences as input.

1. Introduction

Traditionally, most western countries have supplied domestic electricity demand through generation in large central power stations, with subsequent transmission and distribution through networks. The generation efficiency of the power stations varies between around 35% for older coal stations to over 50% for modern combined cycle stations, averaging to about 39%. When transmission and distribution losses are considered, the average overall efficiency of the system drops to 35% [3].

In the coming decade a strong trend towards distributed electricity generation (micro-generation e.g. solar cells, micro Combined Heat and Power (microCHP) appliances, micro gas turbines, micro-windmills, heat exchangers, etc.) is expected.

A microCHP appliance is a system that consumes natural gas and produces heat and — as a by-product during the heat production — electricity. It can generate electricity at the kilowatt level which will allow these units to be installed in an individual home. They can be connected directly to the domestic heating and electrical systems, which leads to a very high efficiency (up to 90%) in usage of primary energy. The heat is used for the heat demand in the home such as central heating, showering, hot water taps etc. The electricity can be used in the home or, when not needed, be exported to the electricity distribution network.

For the stability of the electricity distribution network, it is imperative that production and demand are always in balance. Adding a large number of micro-generators to the grid might disrupt this balance since they are driven by heat-demand (microCHP) or nature (solar cells, micro-windmills), which makes them less controllable.

In case of a microCHP, adding a heat buffer (hot water tank) decouples the demand and production of heat. This gives some flexibility in the electricity production, allowing the production of electricity on more beneficial periods. For example, we may fill the hot water tank when people get home from work during the evening peak. The hot water can be used the next morning for showering, while the produced electricity can be used by the appliances switched on when people get home.

It is expected that microCHP appliances will replace the current high efficiency boilers [6]. This will increase the amount of microCHP appliances on the grid in the near future. When the number of microCHP appliances becomes high enough, generators can be grouped together and become a Virtual Power Plant (VPP). By controlling and smart scheduling such a fleet of generators a virtual power plant may replace a conventional (less-efficient) power plant. Using a virtual power plant instead of a conventional one will result in a significant reduction in costs and CO_2 emission due to a more optimal use of primary energy sources.

Important in such an approach is the controllability of the group of generators. In case of the Dutch electricity market, suppliers (producers) and consumers of electricity have to specify one day in advance what there electricity production/demand is going to be for each quarter of an hour. Every deviation from this specification will result in an imbalance and is penalized by a central watchdog. The deviation has to be compensated elsewhere in the network. As a consequence, to use a virtual power plant, the production capacity of the fleet has to be predicted at reasonable accuracy. This will ensure the promised production capacity is really available.

For microCHP, the electricity production capacity is based on the heat demand. Thus an accurate heat demand prediction is required. In our approach, we predict the heat demand for each individual household using neural network techniques. Goal of our model is to predict the heat profile for the next day as accurately as possible. Since we use the expected heat load to predict how much and at which times we can produce heat (and thus electricity), two criteria are important for our prediction. First, the amount of heat for the day has to be predicted accurately. Secondly, the shape of the expected heat profile has to be determined.

In the following sections, first our approach to the short-term individual heat demand prediction is given. Since we use neural network techniques, a short introduction to the subject is given in Section 3. Details about implementation and the results are given in Section 4 and Section 5 respectively. We conclude this paper with conclusions and future work.

2. Approach

An accurate heat demand prediction is key to determine the production capacity of a virtual power plant of microCHP appliances and is a required input in the scheduling and control algorithms to enable virtual power planting. Although electricity demand prediction is studied quite extensively [2, 1], individual heat demand prediction is an unexplored field. Most demand prediction schemes try to predict the demand for a large area, for example a complete neighborhood.

We want to predict the heat demand for individual households. The goal is to develop a learning system placed in each household, which is able to learn the behavior of the residents. Since this system is installed in individual households, it can gather the necessary local information needed for the predicting. For example, by analyzing the program of the thermostat, holidays can be detected and the prediction can be adjusted accordingly. By accurately predicting the heat demand of individual households, the quality of the total prediction can improve.

A virtual power plant may combine in the order of hundred thousands up to millions of micro-generators. Predicting the heat demand for each group of micro-generators clustered by area requires a lot of computational power. When this is done by a central control system, the system is not scalable. When each household predicts it's own heat demand, the computational power is distributed over the households, which will improve scalability.

Heat demand is mainly influenced by the weather, behavior of the residents and the insulation capacity of the house. It has already been shown that weather information is a relevant input for electricity demand prediction [2]. The insulation of an house is fixed and is unlikely to change very often. For this reason, we do not use this as an input for our model, since we expect the model will learn the characteristics of the home.

The behavior of the residents has a big influence on the heat demand, but it cannot be used in a distinct way as an input. For the prediction, we need to know the behavior one day in advance, which is rather difficult to describe and obtain. Therefore, we try to deduce patterns of the residents based on data from previous days. Furthermore, it is assumed that people have a modern thermostat. Most thermostats are programmed to a fixed schedule. This schedule can be learned by the system.

As input for fitting the parameters of our heat prediction models we use status information of hot water tanks and installed microCHP appliances of four households, kindly made available by Essent and GasTerra (two local energy companies). For each household, the status of the hot water tank and the microCHP appliance were monitored on a minute basis for roughly one year, starting around the beginning of 2007. From this information, the heat demand for these four households is derived by combining the microCHP appliance status and the changes of the tank levels. Since the microCHP units were used for testing, some gaps in the measurements data occurred. All days with less then 1200 of the total 1440 measurements are filtered and not used as input for our model. In our prediction model we use one hour time periods. To get the hourly heat demand is we by sum up the heat demand per 60 minutes.

As weather input for our model, we have used Meteorological Aerodrome Reports (METAR). METAR reports are produced every half hour by weather stations. From weather stations located nearby the households, we extracted temperature information.

As a measure for the behavior of the residents and the characteristics of the house we use the heat demand of former days as input in order to find fixed patterns. However, it is expected that on different days people might have a different living pattern. For this reason, we chose to have different models for each day of the week.

3. Multilayer feed-forward networks

In our approach, we used a multilayer feed forward neural network. Neural networks, as described in [4], are computational models based on biological neurons. They are able to learn, to generalize, or to cluster data. Their operation is based on parallel processing.

A neural network consists of a pool of simple processing units, which communicate by sending signals to each other over a large number of weighted connections. An example of an processing unit, called a neuron, is depicted in Figure 1.

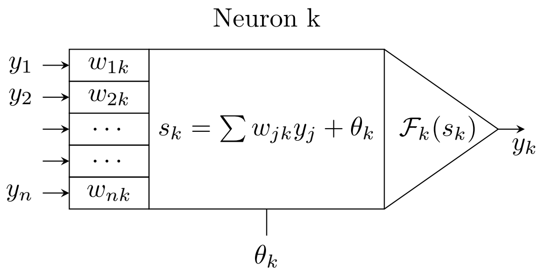

Figure 1. A single processing unit of a neural network

Each neuron basically performs one task. It receives inputs from neighbors and compute an output signal which is propagated to other neurons. Furthermore, in the training phase the neuron also has to adapt the weights of it's input connection to achieve a good fitting to the training data.

Within the neural network three types of neurons exists: input neurons which receive their input from outside the network, output neurons which send data out of the network and hidden neurons whose in- and output remain inside the network.

Neurons are connected to each other via weights w_{jk}, which determines the effect of a signal of neuron j on neuron k. The total input of neuron k normally is simply the weighted sum of the separate outputs of the neurons connected to k plus a bias or offset θ_k, but other propagation rules exists.

The activation function $\mathcal{F}_k$ determines the new level of activation (the output) based on the effective input $s_k(t)$ and current activation $y_k(t)$. In our approach, a sigmoid (S-shaped) function [4] is used as activation function.

A neural network has to be configured (trained) such that the application of the neural network to a set of given input produces the desired outputs (which are also given). The given input/output pairs is the training data. During the training, the weights of the neurons are adapted according to a learning rule. By adjusting the weights, the error between the network output and the expected output is minimized. If a priori knowledge is available, this can be used to pre-specify the weights.

In our approach we use a multi-layer feed-forward (see Figure 2). Each layer consist of neurons which receive their input from a layer directly in front (left in the figure) and send their output to a layer directly behind (right in the figure). There are no connections between neurons within the same layer.

Since we have no a priori knowledge, our model has to be trained completely. When input examples are given to the model, the activation values are propagated to the outputs. Usually there will be an error on the outputs, which has to be minimized (to zero). After determining the error, a backward pass through the network changes all the weights to minimize the error.

4. Implementation

We used MATLAB's Neural Network Toolkit to implement and train our model. When using multi-layer feed-forward networks, different layer sizes and number of layers can be used dependent on the complexity of the system. We have determined the correlation between the heat demand and the heat demand one day earlier, the heat demand one week earlier and the temperature. All three groups are highly correlated ($\rho > 0.85$) with the heat demand. Since there is a high correlation, we chose to use a small amount of layers. In our neural network we use two layers (the input layer is not counted).

We are interested in the heat demand for the upcoming day, given the (expected) outside temperatures and previous heat demand. It is preferable to predict the heat demand as accurate as possible (in the order of minutes). However, the amount of information available to the network is not enough to give such an accurate prediction. For this reason, we will predict the heat demand per hour. In our approach, we use an input vector which consist of three groups of data: (a) the heat demand of the previous day (24 values), (b) the heat demand of the same day one week earlier (24 values) and (c) the average (predicted) temperatures per hour of the day (see Figure 2). The output vector of the model (the last layer) consists of 24 outputs, the expected heat demand for each hour of the day.

The input vectors are normalized between -1 and 1 and separated into three sets: training, validation en test set. To determine the optimal network size for the hidden layer, we have trained networks with one up to twenty hidden neurons for each weekday and for all households. We have used the mean squared error as a measure for the error during training and used the Levenberg-Marquardt method [5] as training function.

All combinations of network sizes, weekday and household have been trained three times to minimize the risk of getting stuck in a local minimum during training.

5. Results

After training the model using the approach given in Section 4, we want to determine the quality of the prediction of the model using the validation set. As mentioned in the introduction, the shape of the heat profile and the amount of heat for the day determine the quality of the prediction. To determine the quality of a prediction, and thus a neural network, we use two errors.

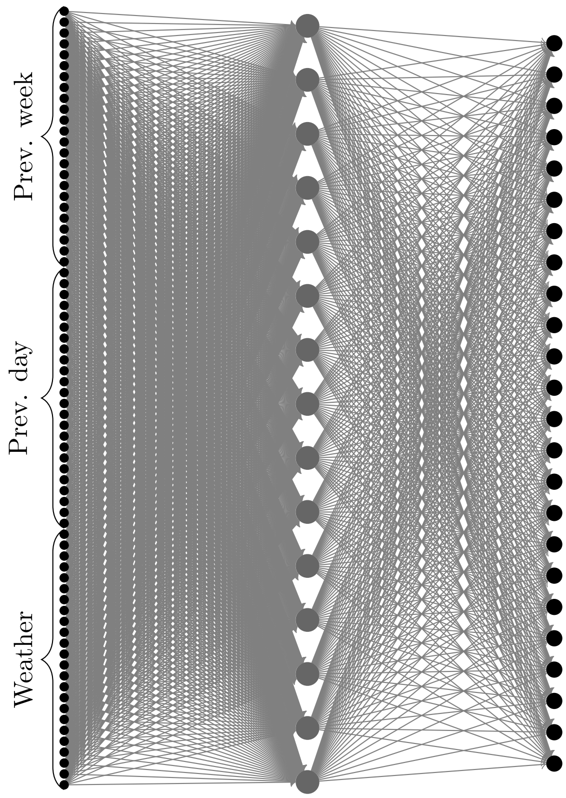

Figure 2. Example of used network structure in the prediction model using 15 hidden neurons.

Table 1. Mean absolute deviation per hour (kWm)

House	Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	41	34	32	43	26	36	24
2	34	34	37	35	43	32	54
3	55	30	54	34	36	43	26
4	45	24	31	27	32	48	37

Table 2. Mean deviation per hour (kWm)

House	Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	13	5	1	-2	6	-15	12
2	5	6	-7	-4	-4	-1	4
3	-20	13	-9	-6	-20	25	1
4	-34	-5	-13	-11	13	-19	-9

Table 3. Mean heat demand (kWm)

House	Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	136	106	110	113	77	118	97
2	151	150	147	167	151	113	205
3	321	179	260	195	286	256	278
4	160	90	103	60	70	112	110

Table 4. Optimal network sizes

House	Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	19	7	5	17	14	6	9
2	7	16	20	14	10	6	12
3	11	12	17	17	9	17	10
4	18	8	19	16	11	6	17

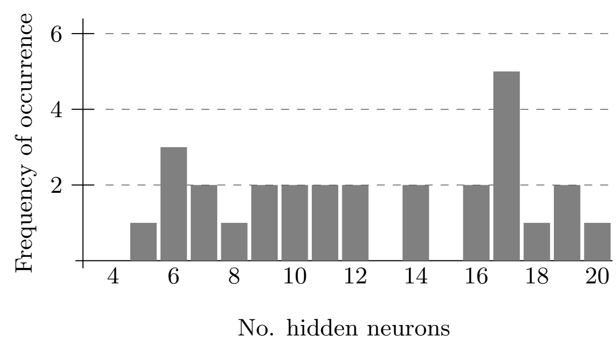

Figure 3. Distribution of optimal network sizes

The first measure is how good the network is able to learn the specific profile of that day. We determine the heat profile for a day by calculating mean heat demand for each hour of the day over a given set. In other words, the mean heat demand between 12 am and 1 am, 1 am and 2 am etc. is determined.

By subtracting the predicted heat profile from the real heat profile, we get a vector of 24 values with the deviation of the profile per hour.

A predicted heat profile should deviate as less as possible from the real profile. For this reason, we determine mean absolute deviation per hour of the best performing models, which is given in Table 1. By using the mean absolute deviation, the total deviation (positive and negative) of the daily profile are measured.

The second measure how the good the network is able to predict the amount of heat per day. Normally, you can sum up the heat demand of the 24 hours of each day to calculate the total heat demand per day and determine the error between the prediction and the real heat demand. However, since we have already determined the deviation per hour, we will use the hourly deviation. By taking the mean error of the hourly deviation, you get the mean deviation per hour. This is the same as determining the total error per day and dividing this sum by 24. In Table 2, the mean deviation per hour is given.

To determine the best performing network, the errors are combined. First, both errors are normalized between 0 and 1, since they have different orders.

The heat profile of the predicted heat demand is more important than the total heat demand for one day. It can be possible that for a day the total heat demand is predicted accurately, but the profile of the prediction is complete off.

In this case, it is possible the predicted production capacity is not available when required. For this reason, we give the profile error a higher a three times higher weight then the day total error (0.75 and 0.25 respectively).

To give an indication of the order of magnitude of the errors, the mean heat demand per hour for each household is given in Table 3. As an example, the heat profile and total demand of household 2 for Saturdays are depicted in Figures 4 and 5 respectively.

If we look at Figure 4, you can see that the global shape of the profile is predicted, but every hour of the day there is an error. This corresponds to the big mean absolute deviation per hour (54 kWm) compared to the mean heat demand per hour (205 kWm).

If we look at Figure 5, you can see that on average, the predicted heat demand is a bit higher than the real demand. This corresponds to a relative small mean deviation per hour (4 kWm) compared to the mean heat demand per hour.

The optimal network sizes for each household and weekday are shown in Table 4. Looking at distribution (Figure 3) of the best performing network size, there is no network size which always perform best for a certain household. However, networks with seventeen hidden neurons in the middle layer is most often the best network size. If we look at the average performance for each network size, networks with sixteen hidden neurons perform best.

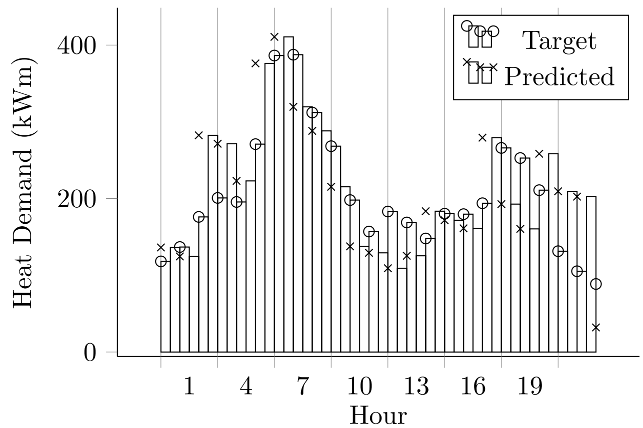

Figure 4. Prediction results of heat profile for household 2, Saturday using 12 neurons

6. Conclusions and future work

We have shown that neural network techniques can be used to predict the heat demand of individual households. When using a network with two layers, using around sixteen to seventeen gives optimal results.

Although there still exists an error, the results are promising. This is our initial design to develop a system which is able to learn the behavior of the residents. Given the limited amount of input, on average the predictions of the heat demand are close to the original heat demands. If we look at the global shape of the heat profile, the trained models show the same global shape.

Because the heat demand is determined by more factors than currently used as input for our models and by human behavior, a certain error will always exist. By adding more parameters to the model, like for example wind speed, the illumination factors and the program of the thermostat, better predictions should be possible. In future models we will include these factors, using the results presented here as a basis.

Furthermore, quite some historical data is used for training. However, it is preferable to use less historical data to make the system more adaptive. If the heat demand changed because of an adjustment in the behavior of the residents or due to seasonal influences, it should not take weeks for the

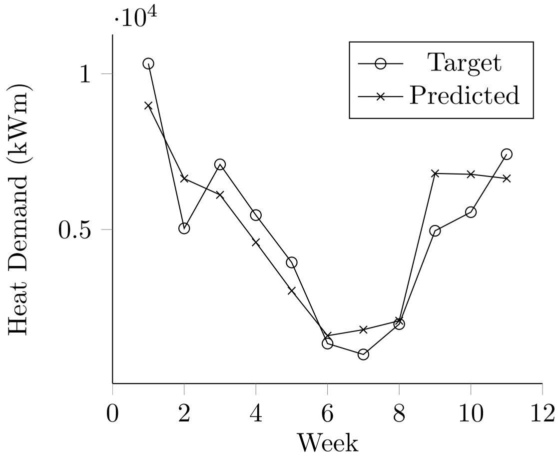

Figure 5. Prediction results of day totals for household 2, Saturday using 12 neurons

system to learn the new demand characteristics. The system has to constantly learn new behavior and might require more logic then the neural networks technique alone.

Finally, we want to predict the heat demand more accurately then in the order of hours.

7. Acknowledgments

This research is funded by Essent and GasTerra.

References

[1] A.-U. Asar and J. McDonald. A specification of neural network applications in the load forecasting problem. *Control Systems Technology, IEEE Transactions on*, 2(2):135–141, Jun 1994.

[2] S.-T. Chen, D. Yu, and A. Moghaddamjo. Weather sensitive short-term load forecasting using nonfully connected artificial neural network. *Power Systems, IEEE Transactions on*, 7(3):1098–1105, Aug 1992.

[3] A. de Jong, E. J. Bakker, J. Dam, and H. van Wolveren. Technisch energie- and CO2-besparingspotentieel van micro-WKK in Nederland (2010-2030). Technical report, Werkgroep Decentraal, July 2006.

[4] B. Krose and P. van der Smagt. An introduction to neural networks, 1993.

[5] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[6] United States Department of Energy. The micro-CHP technologies roadmap. *Results of the Micro-CHP Technologies Roadmap Workshop*, December 2003.